SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
August 15, 1996

3. REPORT TYPE AND DATES COVERED
Final --- 1-23-95 to 6-22-96

4. TITLE AND SUBTITLE
Mobile Testbed for Experiments in Machine Perception and Learning

5. FUNDING NUMBERS
--
DAAH04-95-1-0049

6. AUTHOR(S)
Bir Bhanu

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
University of California
College of Engineering
Riverside, CA 92521

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 34345.1-RT-RIP

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

19960910 028

13. ABSTRACT (Maximum 200 words)
One of the major goals of our research in image understanding is to test and evaluate algorithms under real-world situations. To accomplish this goal, we are developing a mobile platform equipped with sensors and on-board computers.

We use an off-the-shelf electric cart that has been modified and equipped with a pan-tilt camera unit and other hardware for steering and speed control. We have also developed dedicated mechanical and electronic systems for steering and vehicle control, which are monitored by a SUN workstation. Under this Defense University Research Instrumentation grant we have purchased an Imaging Technology Inc. (ITI) real-time image processing hardware, and a SUN workstation. Pan/tilt camera system, steering and speed control, video/sonar sensors and SUN workstation interact with a Z-180 Micro-Controller. Input from video sensors goes to the ITI 150/40 image processing system that is directly connected to the SUN workstation.

(Please see attached continuation page for remainder of item #13.)

14. SUBJECT TERMS
Automated Vision (Sensing Recognition Systems)

15. NUMBER IF PAGES

16. PRICE CODE

DTIC QUALITY INSPECTED 3

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Item 13. Abstract (Continued)

The mobile testbed will be used for advancing the state-of-the-art in algorithms by performing experiments in perception and learning in outdoor environments. Further, the development of this platform has provided much needed hands-on experience to students in system engineering, software, hardware and robust algorithm development.

Mobile Testbed for Experiments in Machine Perception and Learning

Grant No.: DAAH04-95-1-0049

Item	Qty	Description (Part No.)	Source	Contact Person	Cost Unit	Cost Total
Image Processing System						
ITI Image Processing Hardware			Imaging Technology, Inc.	Ron Bryan		
		Model: Series 150/40	2134 Main St., Suite 160	(T) 714/960-7676		
		Vision Processor	Huntington Beach, CA 92648	(F) 714/969-9138		
	1	Image Manager with 4 Mbyte Memory (IMA-VME-4.0-N-H-N)				7,173.00
	1	Color Acquisition Module (AMCLR-H)				1,345.50
	1	Convolver/Arithmetic Logic Unit (CMCLU-H)				3,213.00
	3	Programmable Accelerator (CMPA-H)			3,213.00	9,639.00
	2	Computational Module Controller (CMC-VME-H)			2,313.00	4,626.00
	1	Pseudo Color Display Module (DMPC-H)				963.00
	1	Median Morphological Processor (CMMMP-1-H)				2,763.00
	1	Memory Module, 16 Mbyte (CMMEM-16-H)				5,103.00
	1	Histogram/Feature Extraction Processor (CMHF-H)				2,763.00
	1	ITEX Core Software Source Code for Solaris (ITEX-CORE-VME-SRC-S-SOL-SB)				4,500.00
	1	ITEX CM Software Source for Solaris (ITEX-CM-SRC-S-SOL-SB)				3,375.00
	1	ITEX PA Software Source for Solaris (ITEX-PA-SRC-S-SOL-SB)				4,500.00
	1	Full Development Software for CM-PA Object Code (PDS-PAF-OBJ-S-DOS-P)				4,495.50
	1	Frontplane Video Bus, 5 Slot (FBV-1504-VME-5)				445.50
	1	Breakout Cable, 2 Cameras (BCBL-CAM2)				130.50
	2	Adaptor Cable - Color BNC (ACBL-CLR)			103.50	207.00
	1	S-Bus Translator (BT-15040-SB)				1,975.50
					Sub-Total	57,217.50
					Tax	4,111.65
						61,329.15
	1	MVC 150/40 Installation Package (SR15040-FIC)		*(non-taxable)*		1,995.00
				Image Processing System Total:		63,324.15
Support Equipment			Sun Microsystems	Bill Perkins		
Sun Ultra 1 System			3401 Centrelake Dr., Ste 410	(T) 909/933-5018		
			Ontario, CA 91761	(F) 909/984-1627		
	1	(A12-UBA1-1E-064MA) Model 170E with 167 MHz UltraSPARC Processor				13,717.55
		Creator 24-bit Color Accelerated Graphics and Imaging Workstation			Tax	1,063.11
		(Multimedia Package), 20-inch Color Monitor, Creator Single Buffer, Graphics,				14,780.66
		64 Mbytes, 2.1 Gbyte 5400 RPM Internal Fast/Wide SCSI-2 Disk, Internal SunCD 4				
	1	(A11-UAA1-1A-064AB) Model 140 with 143 MHz UltraSPARC Processor			*(Partial)* *	6,895.19
		TurboGX 8-bit Accelerated Graphics Workstation, 20-inch Color Monitor,				
		Turbo GX 1-Mbyte Frame Buffer, 64 Mbytes, 2.1 Gbyte, 5400 RPM Internal Fast SCSI-2Disk				
				Support Equipment Total:		21,675.85
				TOTAL		**85,000.00**

* Partial of item paid.

1 The Mobile Testbed

1.1 Purpose and Goals

The purpose of our mobile testbed is to support research and experimentation in the areas of computer vision, sensor-based robot control, autonomous navigation, and applied machine learning in outdoor scenarios under realistic environmental conditions. The testbed carries a range of different sensors, special-purpose hardware for sensor data processing, and general-purpose computers to perform these tasks. The key design goals are flexibility and processing speed. Flexibility is important in order to accommodate and evaluate different processing and control paradigms. Processing speed must be sufficient for the system to perform under real-time (or near real-time) constraints.

In this section we describe our testbed, its status and the student participation. A summary of our current research and the state-of-the-art in perception-based navigation is included in sections 2 and 3, respectively.

1.2 The Mobile Platform

The platform (see Figure 1) we are using is an off-the-shelf electric cart manufactured by Taylor-Dunn in Anaheim, California. It has a three-wheel chassis with two driven rear wheels and a single steerable front wheel, which provides a simple steering geometry. The cart is driven by a DC motor powered by 24V lead batteries, which also powers all on-board equipment. The cart was customized by the manufacturer with an extended steering shaft that connects to the electric steering mechanism.

During the operation of the vehicle, a human operator is required to ride on the vehicle for safety reasons. The operator has full mechanical override capabilities with respect to steering and braking.

1.3 Sensors

The test platform is equipped with the following sensors:

1. Two *color video cameras* in a binocular stereo configuration, mounted on a pan-tilt head. To keep the weight of movable parts low, cameras with detached sensor heads are used.

   For the color video cameras we use Sony XC-711RR. This is a single-chip line-transfer CCD camera with a detached sensor head, which is important to keep the weight of

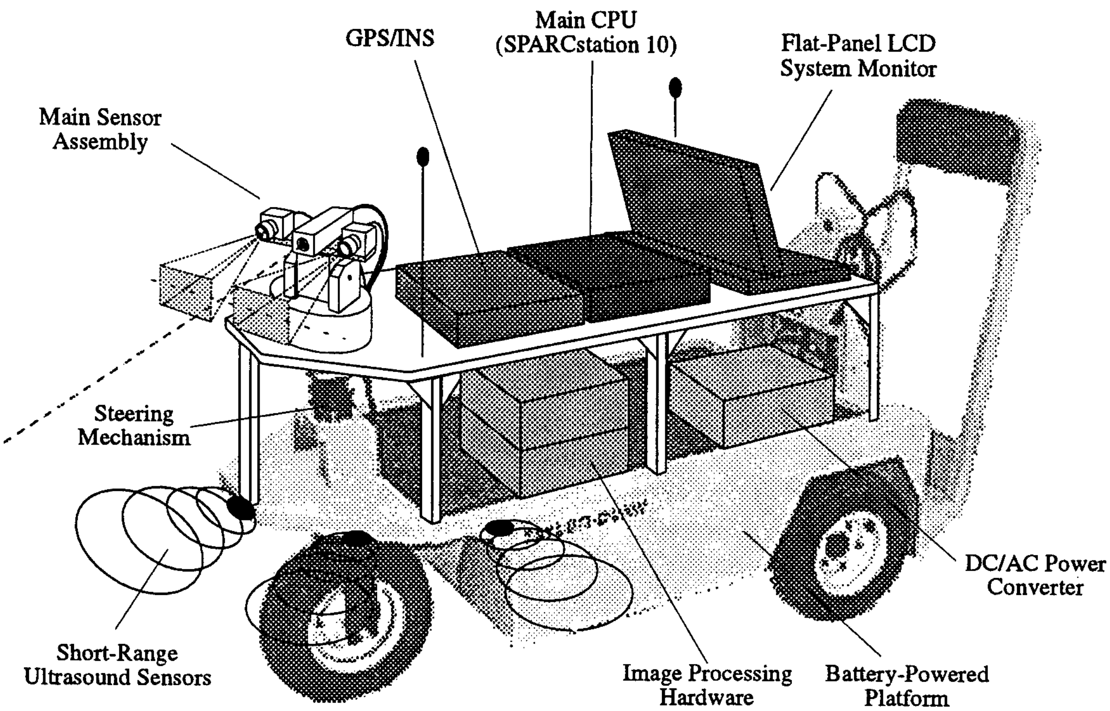

Figure 1: Mobile testbed for machine perception and learning research. Built on a battery-powered three-wheel platform, the testbed carries a range of sensors and on-board computers for real-world experiments.

the moving parts low. The camera unit produces standard NTSC video signals in interlaced or non-interlaced mode. The cameras are equipped with a pair of fixed-focus wide-angle lenses which provide a large viewing angle, large depth of focus, and are relatively simple to calibrate.

2. A single *Laser point ranging system*, mounted on the same pan-tilt head as the two video cameras (Figure 2).

   The main use of the Laser range finder is to provide direct distance measurements on objects too far away or with insufficient texture for stereo measurements, thus supporting surface interpolation and calibration. The proposed Laser range finder (High-Accuracy Altitude Measurement System, manufactured by Schwartz Electro-Optics, Inc.) is a single-point ranging device based on a solid-state infrared Laser source and time-of-flight measurement. The operating range with non-cooperative targets is between 3—100 meters, with 0.1 meter accuracy (the range for cooperative targets is up to 300 meters). The unit is eye safe. The entire range finder system is contained in a compact housing similar to a soda can in both shape and size. It weighs approximately 2 kg. Range measurements are communicated to the main on-board processor through serial (RS-232) links. Due to the unavailability of funds we do not have a laser range finder at this time.

3. A set of ultrasound sensors for close-range maneuvering and collision avoidance. We have purchased and tested Polaroid ultrasound sensors (newer product). They have yet to be installed on the vehicle.

In addition, the platform will carry the following direct state sensors:

1. A differential GPS receiver to support map-based navigation and environmental exploration.

2. An inertial navigation system for continuously monitoring vehicle motion.

2 Image Processing Hardware

The on-board image processing hardware consists of a multi-board VME-bus system that is directly connected to the SUN Ultra 1 computer via a SBus-to-VME interface. The key selection criteria for this hardware were flexibility, performance, operating system support (Solaris), software support, and cost. In addition to common image processing requirements

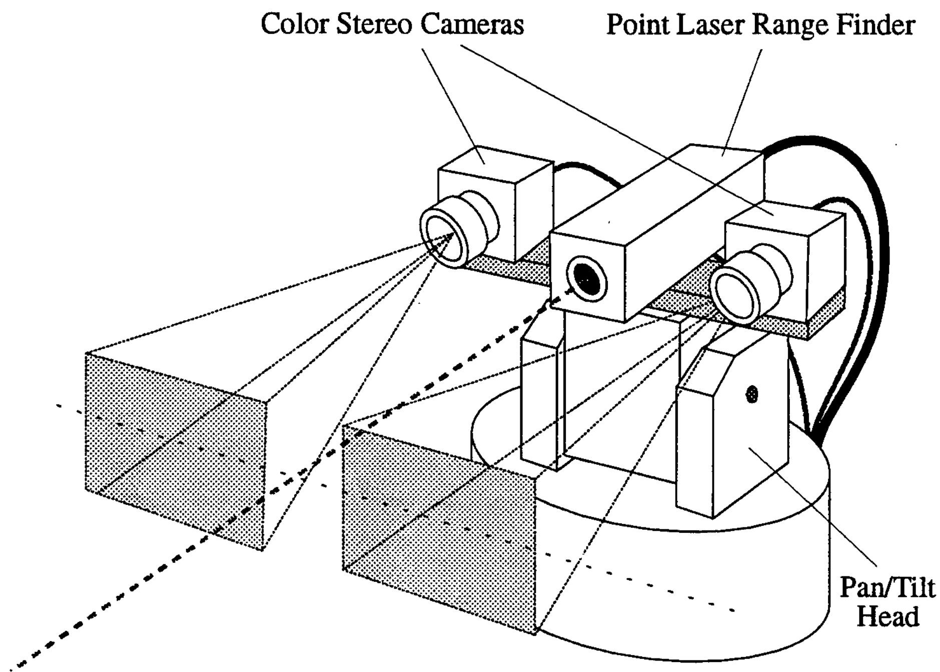

Figure 2: Main sensor assembly. Two color video cameras and a point Laser range finder are mounted on a common pan-tilt mechanism. The two cameras are arranged in a binocular stereo configuration with fixed vergence. The orientation of the Laser range finder with respect to the stereo cameras is fixed. The Laser ranger can be pointed at selected scene entities for depth verification by pan and tilt motion.

(fast convolution, point operations, etc.), we also want to have the possibility to use general-purpose CPUs with floating-point capabilities and fast access to the image data (e.g., the TI 320Cxx or Intel 860 processor series). We narrowed down our evaluation of available systems to (a) the MaxVideo 200 from Datacube and (b) the Series 150/40 from Imaging Technology. During this evaluation, comments from current users of both systems that we solicited through the Internet were of great help. Based on the results of this evaluation, we have specified an ITI Series 150/40 Vision Processor. The hardware architecture is shown in Figure 3.

2.1 Student Participation and Current Status

Both graduate and undergraduate students have participated in the development of the testbed. Six Electrical Engineering Senior students (Mardi Ouch, Mike Miles, William Saw, Brian Schroeder, Diane Heck and Tony Sterling) developed the vehicle during their two quarter senior project. They developed the necessary hardware and software system that runs on the SUN Ultra 1 computer and controls the steering, speed, pan-tilt mechanisms of the vehicle. Songnian Rong, a graduate student in Computer Science, helped in specifying motors, gears, materials, power supplies, transformers, etc. Students have carried out electrical, mechanical, graphic interface, and software aspects of the project quite successfully. Both the propulsion (braking) and the steering are controlled from the main on-board processor (Sun Ultra 1 Computer) through serial (RS-232) communication links. The steering mechanism consists of a stepper motor and a planetary gear with a 100:1 ratio. A dedicated controller hardware has been built for the propulsion motor. We have evaluated the image processing system. During the coming academic year we will develop the tracking, gaze control, automated landmark acquisition and path retrace capabilities using the vehicle.

3 Current Research in Multistrategy Learning for Image Understanding

Current work in this area at UCR is focused on applying multiple machine learning strategies for solving fundamental problems and improving performance in Image Understanding. Major support for this work is provided through a grant by ARPA/AFOSR (F49620-95-1-0424) and results have been reported in DARPA IU Workshops 1994 and 1996. The application of machine Learning (ML) to the Image Understanding (IU) domain is more demanding than most conventional learning applications in AI. It is caused by (a) the enormous amount of incoming data to be processed, and (b) the variety of processes and representations encountered in Image Understanding.

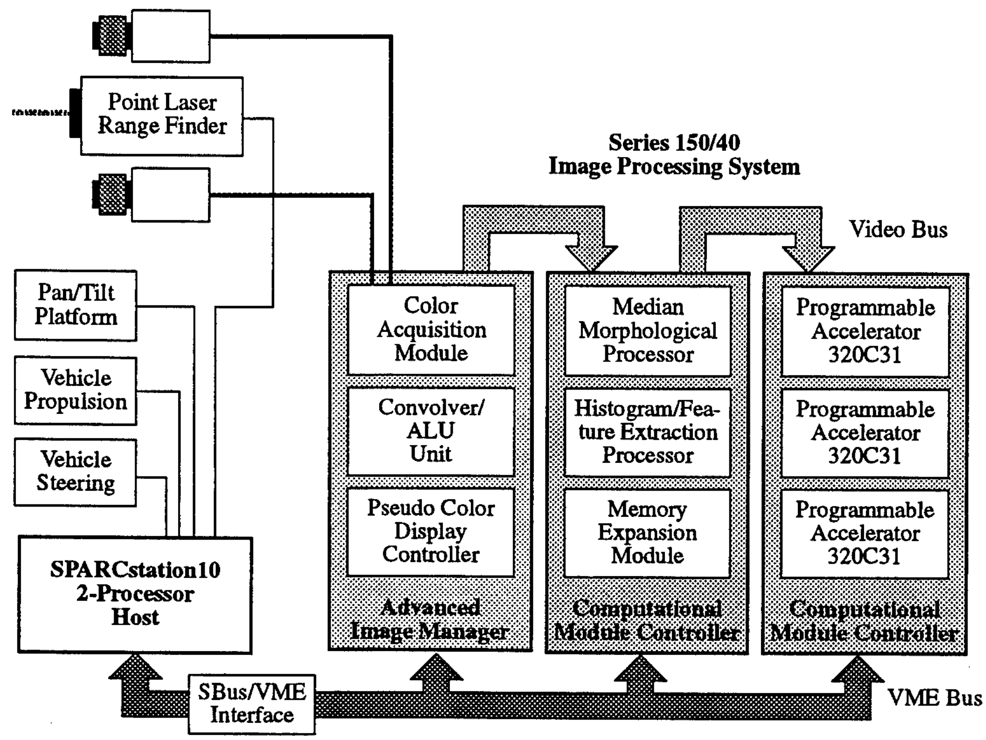

Figure 3: Hardware architecture for the mobile testbed. The main host is a dual-processor Sun Ultra computer, which is connected to the ITI Series 150/40 Image Processor through a SBus/VME bus interface. The image processing system consists of (a) an Advanced Image Manager (IMA15040-4-V) with a Color Acquisition Module (AMCL), a Convolver/ALU Unit (CMCLU), and a Pseudo Color Display Controller (DMPC), (b) two Computational Module Controller (CMC15040-V) with a Median Morphological Processor (CMMP-1), a Histogram/Feature Extraction Processor (CMHF), a Memory Expansion Module (CMMEM), and three Programmable Accelerators (CMPA). The Sun computer controls the vehicle steering mechanism, vehicle propulsion, the pan/tilt sensor platform, and the point laser range finder through serial (RS-232) communication lines. Note that the SPARCStation 10 has been replaced by Sun Ultra 1 Computer

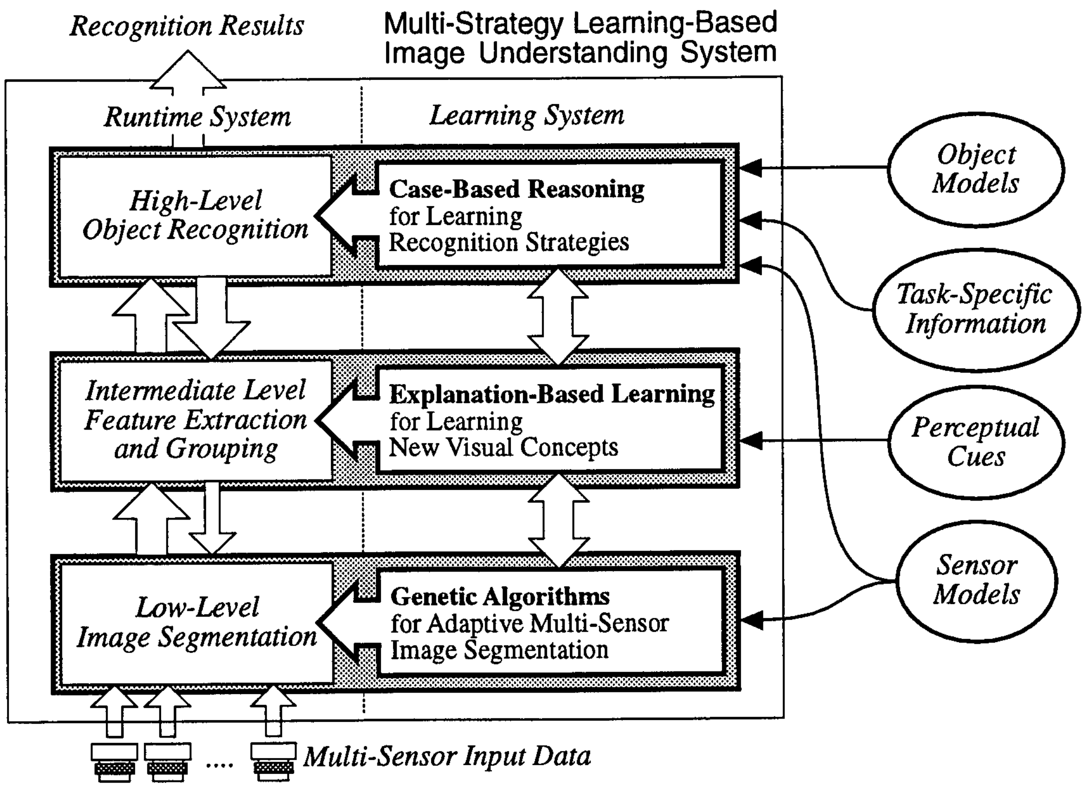

Figure 4: Multistrategy Learning for Image Understanding [5].

Key Ideas

The multi-strategy learning-based IU System (Figure 4) selectively applies machine learning techniques at multiple levels to achieve robust recognition performance. The system uses Genetic Algorithms (GAs) to optimize multi-sensor image segmentation at the low level. At the intermediate level, Explanation-Based Learning (EBL) is employed to learn new visual concepts for improving indexing and matching. At the high-level, Case-Based Reasoning (CBR) is used to dynamically adapt recognition strategies, and acquiring and maintaining information about the environment. At each level, appropriate evaluation criteria are employed to monitor the performance and self-improvement of the system. We also use Hidden Markov Models (HMM) for signal-to-symbol conversion and reinforcement learning for feedback between different levels. Our goal is to demonstrate new learning techniques in the context of navigation and automatic target recognition problems.

4 State-of-the-Art in Perception-Based Navigation

The field of vision-based navigation has continuously advanced from the early algorithms for road following and limited cross country navigation for land vehicle navigation to developing systems for more complex environments, to detecting obstacles like trees, wires, etc., for helicopter navigation, to reconstructing 3-D models for small underwater objects. While the requirements and associated design details of the vehicles for land, air and underwater navigation are different, many of the underlying problems and the scientific principles used for sensing, processing, and implementation of perception-based robotic systems in these environments have some common characteristics. The state of the field has been summarized in a survey paper on perception-based outdoor navigation [13].

5.1 Machine Perception

The specific tasks to be performed by a perception-based navigation system strongly depend upon the particular mission and application domain. (Figure 5).

Localization

Localization is the problem of determining the current position and orientation of the vehicle with respect to a given map. Perception-based localization amounts to specifying the current viewpoint in a world coordinate system and matching the sensor data to the expected view obtained from the map data that consist of Digital Terrain Elevation Data (DTED) and Digital Feature Analysis Data (DFAD). The enormous search space of possible viewpoints, combined with the highly non-linear effects of changing viewpoints on the sensed data make the problem quite difficult [22, 43, 51]. Naturally, the problem is simplified when the approximate vehicle location, orientation, and elevation are known, or when the scenes contain distinctive elements (landmarks).

Object Recognition

Object recognition addresses the general problem of identifying objects in a scene. In the context of autonomous navigation, object recognition is important because certain decisions depend upon the semantic categories of the observed objects. Object recognition goes beyond both obstacle and landmark recognition. The representation used for object recognition are usually more specific than those for obstacle detection and more general than those used in landmark recognition. In the case of highway navigation, important

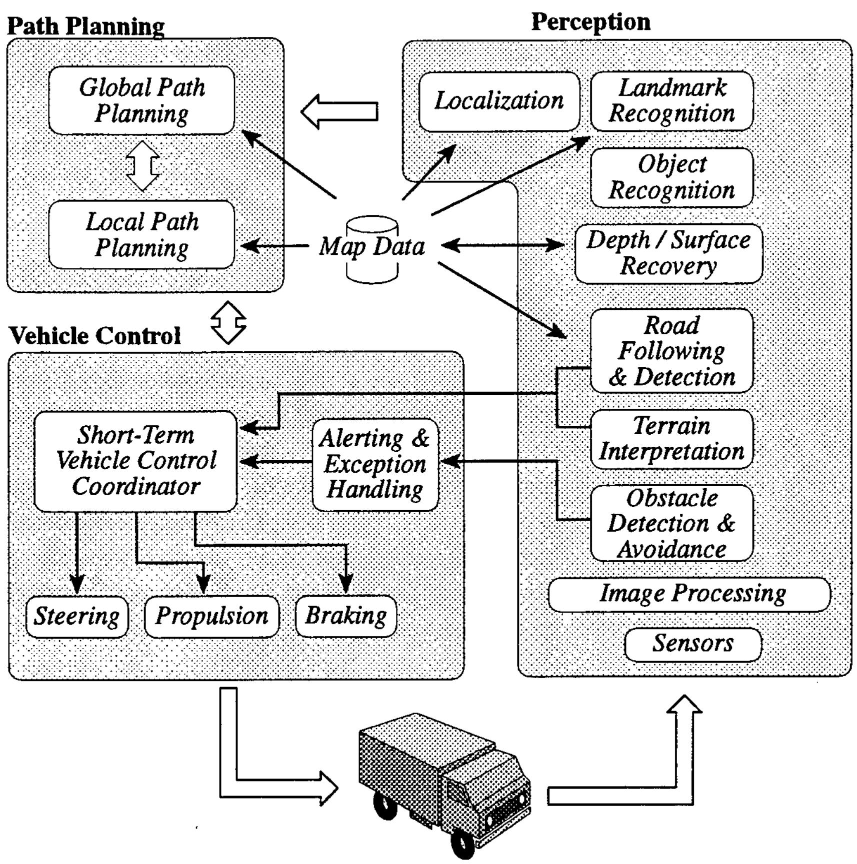

Figure 5: Typical tasks and subtasks involved in a perception-based land navigation system include localization, landmark recognition, object recognition, depth and surface recovery, road following, terrain interpretation, obstacle detection and avoidance.

landmarks, such as traffic signs, signals, ground markings etc., are always placed in obvious locations, which makes the task more tractable than general object recognition. The main problem in outdoor object recognition is that, although the number of relevant object categories may not be very large, the appearance of the objects encountered may vary significantly between individual instances. This is not only true for natural objects (e.g., trees) but also for many man-made objects, such as cars and buildings. Also, objects may appear in a variety of different views, unlike in landmark recognition, where the viewpoint is often highly constrained. The semantic interpretation of complex traffic scenes, involving several moving cars, pedestrians, and possibly other objects, is a challenging problem made difficult by incomplete evidence, e.g., due to partial occlusion of objects.

Depth and Surface Recovery

In the context of navigation, the availability of depth information is crucial for path planning, obstacle detection and avoidance, exploration, and geographical data acquisition. Depth and surface reconstruction can be based on active or passive sensing techniques. Active sensors, such a laser range finders (Ladar), millimeter-wave Radar, and acoustic devices, provide direct range measurements in either a scanning or selective-focusing mode. Passive range estimation techniques include stereo vision [4, 16, 22, 33, 35], structure-from-motion techniques [11, 25], and combinations thereof [12, 24, 41]. However, current techniques still need significant improvements for successful navigation based on passive ranging [10]. While passive ranging techniques can potentially supply a dense field of range measurements, they rely upon sufficient image texture in the area of interest. Since active sensors do not have this limitation, complementary use of active and passive sensing techniques can improve the reliability and coverage of a depth and surface reconstruction system.

Road Following and Detection

Road detection can be classified as a special case of terrain interpretation. Roads are constructed following certain conventions which can be exploited to improve detection. Highways and major roadways are usually well-marked with lane and edge markers that have contrasting colors compared to the road surface. In such cases, the markers can be extracted as line segments or points denoting the lane boundary. Roads typically have approximately constant width and limited local curvature. These properties can be used in a road model to integrate local measurements into a scene interpretation that is robust to misclassification.

Many road following and detection algorithms have been developed at various institutions [18, 45, 39, 27, 20] The key challenge in road detection is to obtain real-time detection and

to exploit the human conventions of road design while being robust to these non-ideal phenomena.

Terrain Interpretation

Terrain interpretation attempts to characterize the environment with respect to some goal. For the purpose of vehicle navigation, the terrain is mainly characterized by its traversability, which depends upon such factors as the type of the vehicle, the intended speed, direction of traversal, and weather conditions. Terrain interpretation can also be used as the goal of autonomous navigation, in the sense of terrain exploration and map data acquisition. An example for terrain interpretation in the context of ALV cross-country navigation is described in [8], where interpretation is based on multi-spectral image data and contextual constraints.

Obstacle Detection and Avoidance

Dealing with obstacles thus consists of two main tasks: *(a)* the detection and characterization of obstacles and *(b)* performing actions to avoid the obstacles. Obstacle *detection* is sensory-based localization of objects that could impair the planned actions of the vehicle. This includes stationary objects not known *a priori* (e.g., rocks), stationary objects that may change their properties over time (e.g., bushes and potholes), as well as any moving objects. Obstacle *avoidance* attempts to maneuver the vehicle to avoid contact with detected obstacles. Obstacle avoidance can be decomposed into two classes. Within the first class, sensed objects are combined into a local map and a planner chooses a suitable path. The other type of obstacle avoidance is a reflexive action of the vehicle to the sudden, and unexpected, presence of an obstacle. The decision may be qualitative, such as "steer right" or "brake".

Most perception-based approaches for obstacle detection are designed for operation in a stationary environment [3, 33, 36, 46], where the range of an object and its time-to-collision with the vehicle are equally valid measures of an object's proximity. As a result, most research has concentrated on localizing the 3-D position of objects, and maintaining the positional representation as the vehicle moves [3, 36]. In a *dynamic* environment, object range is no longer the best measure of proximity; time-to-collision becomes more important because it accounts for any object motion.

Parallel Processing and Hardware Systems

Many low-level signal and image processing tasks, such as convolution, segmentation, labeling, Hough transforms, and pyramid algorithms, exhibit a high degree of spatial and temporal regularity that make them well-suited for implementation on parallel architectures [7, 48]. In addition, there has been considerable effort to implement high-level processing tasks, such as road following and object recognition, on parallel machines. Some of these machines are described in the following, including the Warp, the iWarp, the Connection Machine, the MasPar systems, and the Image Understanding Architecture.

The *Warp* computer [50], designed at CMU and built by General Electric, has been employed in the CMU Navlab project and in the Autonomous Land Vehicle (ALV) project sponsored by ARPA. The Warp consists of a one-dimensional systolic array of 10–20 processors, each of which can perform 10 million floating-point operations per second. In CMU's Navlab environment, Warp has been used for a variety of tasks, including feature-based stereo (FIDO), collision avoidance based on range images, color-based road following (SCARF), as well as neural network-based road following (ALVINN) [18].

The *iWarp* ("integrated warp") developed by CMU and Intel is a parallel, distributed memory architecture that efficiently supports various inter-processor communication styles, including message passing and systolic communication [23, 37]. Memory communication is flexible and intended for general computing, whereas systolic communication is efficient and well suited for speed-critical applications. CMU has developed high-level software support and computer vision applications for the iWarp [49].

The *Connection Machine* (CM) built by Thinking Machines is a massively parallel SIMD with processors arranged in a hypercube topology [26]. A variety of perception-related algorithms have been implemented on this hardware, including image segmentation, stereo matching, pyramid algorithms, object recognition, and planning [19, 30, 40, 44]. The MIT Vision Machine [38] is an extensive software environment that has been built around the CM.

The *MasPar* is a general-purpose SIMD computer system with scalable architecture based on a reduced instruction set (RISC) design [9]. Its architecture provides not only high computational capability, but also a mesh and global interconnect style of communication. It achieves peak computation rates beyond a billion floating point operations per second. The MasPar MP-1 and MP-2 have been used for road following in CMU's Navlab project [27].

The *Image Understanding Architecture* (IUA), which is being developed at the University of Massachusetts and Hughes, is a multi-level hardware architecture that incorporates several different forms of parallel computation [21]. The lowest level of the IUA consist of

a 512×512 array of 1-bit SIMD processors, that directly operate on the incoming sensor data. A 64×64 array of 16-bit processors mainly performs grouping operations at the intermediate level, running in either SIMD or MIMD mode. High-level, knowledge-based processing, using Lisp and blackboard structures, is performed on 64 coarse-grained RISC processors. Construction of the IUA has been accompanied by a significant software effort, which includes dedicated programming languages and language extensions, libraries, simulators, and graphical user interfaces.

System Architecture

Autonomous navigation systems involve a great variety of computational tasks at different abstraction levels and with varying communications requirements. Depending on the nature of these tasks, different modules may be implemented in radically different ways. For example, low-level image processing tasks are often executed on special purpose hardware, using assembly language or optimized C code. On the other extreme, a high-level planning module may rely on dynamic data structures, knowledge representations, and inference engines that are traditionally implemented in Lisp-like programming languages. As a consequence, real-world navigation architectures are complex, heterogeneous systems and constitute a considerable software (and hardware) engineering challenge. In addition, these systems must be fast, reliable, and should degrade gracefully in case of malfunction. Time constraints are usually a critical factor in perception-based navigation, affecting algorithm design, data communication, and operating system considerations.

Many of the issues addressed in the context of sensor-based robot control architectures apply to perception-based navigation systems as well. In particular, concepts of layered control, such as the subsumption architecture [14], and hierarchical control have been very influential. For example, the Real-Time Control System (RCS) [1] is a hierarchical architecture for intelligent control systems that has been used on the UGV robotics testbed vehicle. NASA and NIST have developed the Standard Reference Model (NASREM) [31], a hierarchical architecture for sensor-based telerobot control systems. NASREM is intended to provide a flexible testbed for research in perception-based robotics, mainly for space applications. PROTEUS [28] is a hybrid between a highly structured hierarchical control system (such as NASREM) and a purely distributed layered control system (such as the subsumption architecture). For land navigation, a large number of experimental system architectures have been implemented at various institutions, such as the Navlab architecture at CMU [42, 47] and the Hughes ALV navigation architecture [36].

5.2 Adaptation and Learning

Adaptive systems are capable of dynamically adjusting their behavior to the current situation and can therefore achieve better performance than static systems. In contrast to *adaptation*, *learning* incorporates memory that is invoked to cope with situations that have been encountered sometime in the past. The need for adaptation and learning capabilities in the context of autonomous navigation arises mainly in the areas of actuator and sensor control, data processing, and map and landmark acquisition. which are discussed below.

Learning for Actuator and Sensor Control

A maneuverable vehicle usually incorporates several actuators that must be operated in a coordinated fashion for proper navigation. Learning can be used to obtain basic input/output relationships, which are often difficult to model explicitly off-line, as well as to refine the existing actuator behavior [15, 17, 32, 39]. With active, goal-directed sensing, the problem of coordinating maneuvers and sensor control adds additional tasks that can be supported by learning, such as gaze stabilization and object tracking.

Learning for Sensor Data Processing

Current machine perception techniques lack the required robustness, reliability, and flexibility to cope with the large variety of situations encountered in a real-world environment. Many existing techniques are brittle in the sense that even minor changes in the expected task environment (e.g., different lighting conditions, geometrical distortions, changing vegetation, etc.) can strongly degrade the performance of the system or even make it fail completely [10]. The introduction of adaptive and learning techniques to sensor data processing has been identified as a key ingredient for building more robust perception systems [34]. Applications include (a) low-level processing, such as adaptive image segmentation, (b) feature selection and grouping at the intermediate level, and (c) high-level tasks, such as 3-D object recognition [6].

Map and Landmark Acquisition

Map information is important for many navigation task, in particular for path planning. The problem of map acquisition arises either when the vehicle navigates through an unknown environment for which no map is yet available, or when the given map data are not sufficiently accurate and need to be updated. Maps for autonomous navigation are often

adapted to the given task and processing environment and may, therefore, differ substantially from the annotated 2-D maps used in everyday life. For example, these maps may consist of a hierarchy of local and global map data that facilitate more efficient navigation [2].

The acquisition of landmarks involves detecting and classifying suitable landmark objects, evaluating their properties, and positioning them in a global map. Landmarks should be distinctive and easy to detect with the given sensors. Models of landmarks can be obtained in variety of ways, e.g., they can be supplied manually, derived from existing data (e.g., maps, digital elevation models, CAD data), or acquired at the actual location.

A major difficulty involved in map data acquisition is the estimation and representation of measurement errors. Also, the updating mechanism must be able to successively reduce this uncertainty by incorporating additional sensor information, either from separate observations or data from multiple sensors. Mapping schemes used for *exploration* must also be able to represent unexplored regions or unknown views of map objects [29].

6 Bibliography

[1] J.S. Albus. RCS: a reference model architecture for intelligent control. *Computer*, 25(5):56–59, May 1992.

[2] M. Asada. Map building for a mobile robot from sensory data. *IEEE Trans. Systems, Man, and Cybernetics*, 37(6):1326–1336, 1990.

[3] N. Ayache and O.D. Faugeras. Maintaining representations of the environment of a mobile robot. *IEEE Journal of Robotics and Automation*, 5(6):804–819, 1989.

[4] S.T. Barnard. Stochastic stereo matching over scale. *International Journal of Computer Vision*, 3:17–32, 1989.

[5] B. Bhanu, W. Burger, S. Das, S. Rong, and X. Wu. Multistrategy learning for image understanding. Annual Report VISLAB-MSL-93-1, ARPA/AFORS Grant Number F49620-93-1-0624, University of California, Riverside, CA, February 1994.

[6] B. Bhanu and S. Das. *Computational Learning for Adaptive Computer Vision.* Plenum Publishing Company, 1996.

[7] B. Bhanu, B. Hutchings, and K. Smith. VLSI design and implementation of a real-time image segmentation processor. *International Journal of Machine Vision and Applications*, 3(1):21–44, January 1990.

[8] B. Bhanu and P. Symosek. Interpretation of terrain using hierarchical symbolic grouping for multi-spectral images. In *Proc. DARPA Image Understanding Workshop*, pages 466–474, 1987.

[9] T. Blank. The MasPar MP-1 architecture. In *Proc. IEEE Computer Society Int. Conf. (COMPCON)*, pages 20–24, San Francisco, 1990.

[10] R.C. Bolles, H.H. Baker, and M.T. Hannah. The JISCT stereo evaluation. In *Proc. DARPA Image Understanding Workshop*, pages 263–274, April 1993.

[11] R.C. Bolles, H.H. Baker, and D.H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1:7–55, 1987.

[12] R.N. Braithwaite and M.P. Beddoes. Obstacle detection for an autonomous vehicle. In *Proc. Conf. on Military Robotic Applications*, pages 206–213, Medicine Hat, Alberta, Canada, September 1991.

[13] R.N. Braithwaite, W. Burger, and B. Bhanu. Perception for outdoor navigation. *IEEE Transactions on Robotics and Automation.* under revisions.

[14] R.A. Brooks. A robust layered control system for a mobile robot. *IEEE Trans. Robotics and Automation*, 2(1):14–23, March 1986.

[15] R.A. Brooks and M.J. Mataric. Real robots, real learning problems. In J.H. Connell and S. Mahadevan, editors, *Robot Learning*, chapter 8, pages 193–213. Kluwer Academic Publishers, 1993.

[16] S.D. Cochran and G. Medioni. Accurate surface description from binocular stereo. In *Proc. DARPA Image Understanding Workshop*, pages 857–869, 1989.

[17] J.R. Cooperstock and E.E. Milios. Self-supervised learning for docking and target reaching. *Robotics and Autonomous Systems*, 11:243–260, 1993.

[18] J.D. Crisman and C.E. Thorpe. SCARF: A color vision system that tracks roads and intersections. *IEEE Trans. Robotics and Automation*, 9:49–58, 1993.

[19] L.S. Davis, D. DeMenthon, T. Bestul, and D. Harwood. RAMBO-vision and planning on the Connection Machine. In *Proc. DARPA Image Understanding Workshop*, pages 631–639, 1989.

[20] E.D. Dickmanns, B. Mysliwetz, and T. Christians. An integrated spatio-temporal approach to automatic visual guidance of autonomous vehicles. *IEEE Trans. Systems, Man, and Cybernetics*, 20(6):1273–1284, 1990.

[21] C. Weems et al. Status and current research in the Image Understanding Architecture program. In *Proc. DARPA Image Understanding Workshop*, pages 1133–1140, 1993.

[22] D.B. Gennery. Visual terrain matching for a Mars rover. In *Conf. on Computer Vision and Pattern Recognition*, pages 483–491, June 1989.

[23] T. Gross. An overview of programming the iWarp system. *International Journal of High Speed Computing*, 5(3):379–401, September 1993.

[24] C. Harris, M. Stephens, E. Sparks, and M. Pike. DROID analysis of the NASA helicopter images. *Journal of Robotic Systems*, 9(6):773–785, September 1992.

[25] D.J. Heeger. Optical flow from spatiotemporal filters. In *Proc. International Conf. on Computer Vision*, pages 96–100, 1987.

[26] W.D. Hillis and L.W. Tucker. The CM-5 Connection Machine: A scalable supercomputer. *Communications of the ACM*, 36(11):30–40, 1993.

[27] T.M. Jochem, D.A. Pomerleau, and C.E. Thorpe. MANIAC: A next generation neurally based autonomous road follower. In *Proc. DARPA Image Understanding Workshop*, pages 473–473, 1993.

[28] M. Kao, G. Weitzel, X. Zheng, and M. Black. A simple approach to planning and executing complex AUV missions. In *Proc. Symposium on Autonomous Underwater Vehicle Technology*, pages 95–102, Washington, DC, June 1992.

[29] R. Kumar and A.R. Hanson. Model extension and refinement using pose recovery. *Journal of Robotic Systems*, 9(6):753–772, September 1992.

[30] J.J. Little, G. Blelloch, and T. Cass. Parallel algorithms for computer vision on the Connection Machine. In *Proc. DARPA Image Understanding Workshop*, pages 628–638, 1987.

[31] R. Lumia, J. Fiala, and A. Wavering. The NASREM robot control system and testbed. *International Journal of Robotics and Automation*, 5(1):20–26, 1990.

[32] S. Mahadevan and J.H. Connell. Automatic programming of behavior-based robots using reinforcement learning. *Artificial Intelligence*, 55:311–365, 1992.

[33] L. Matthies. Passive stereo range imaging for semi-autonomous land navigation. *Journal of Robotic Systems*, 9(6):787–816, 1992.

[34] S. Neghdaripour and A. Jain, editors. *Final Report of the NSF Workshop on the Challenges in Computer Vision Research; Future Directions of Research*, Washington, D.C., November 1991.

[35] H.K. Nishihara. Practical real-time imaging stereo matcher. *Optical Engineering*, 23(5):536–545, October 1984.

[36] K.E. Olin and D.Y. Tseng. Autonomous cross-country navigation: an integrated perception and planning system. *IEEE Expert*, pages 16–30, August 1991.

[37] C. Peterson, J. Sutton, and P. Wiley. iWarp: a 100-MOPS, LIW microprocessor for multicomputers. *IEEE Micro*, 11(3):81–87, June 1991.

[38] T. Poggio, J. Little, E. Gamble, W. Gillett, D. Geiger, D. Weinshall, M. Villalba, N. Larson, T.Cass, H. Bulthoff, M. Drumheller, P. Oppenheiner, W. Yang, and A. Hurlbert. The MIT Vision Machine. In *Proc. DARPA Image Understanding Workshop*, pages 177–198, 1988.

[39] D.A. Pomerleau. Progress in neural network-based vision for autonomous robot driving. In I. Masaki, editor, *Proc. Intelligent Vehicles*, pages 391–396, 1992.

[40] I. Rigoutsos and R. Hummel. Massively parallel model matching – geometric hashing on the Connection Machine. *Computer*, 25(2):33–42, 1992.

[41] B. Sridhar and R. Suorsa. Comparison of motion and stereo in passive ranging systems. *IEEE Trans. Aerospace and Electronic Systems*, 27:741–746, July 1991.

[42] A. Stentz. The CODGER system for mobile robot navigation. In C.E. Thorpe, editor, *Vision and Navigation: the Carnegie Mellon Navlab*, chapter 9, pages 187–201. Kluwer Academic Publishers, 1990.

[43] R. Talluri and J.K. Aggarwal. Position estimation for an autonomous mobile robot in an outdoor environment. *IEEE Trans. on Robotics and Automation*, 8:573–584, 1992.

[44] D.W. Thompson and J.L. Mundy. Model-directed object recognition on the Connection Machine. In *Proc. DARPA Image Understanding Workshop*, pages 98–106, 1987.

[45] C. Thorpe, M. Hebert, T. Kanade, and S. Shafer. Toward autonomous driving: the CMU Navlab, Part I – Perception. *IEEE Expert*, 6(4):31–42, August 1991.

[46] C. Thorpe, M.H. Hebert, T. Kanade, and S.A. Shafer. Vision and navigation for the Carnegie-Mellon Navlab. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(3):362–373, 1988.

[47] C.E. Thorpe, editor. *Vision and Navigation: the Carnegie Mellon Navlab*. Kluwer Academic Publishers, 1990.

[48] G.S. Van Der Wal and P.J. Burt. A VLSI pyramid chip for multiresolution image analysis. *International Journal of Computer Vision*, 8:177–189, 1992.

[49] J.A. Webb. Implementation and performance of fast parallel multi-baseline stereo vision. In *Proc. DARPA Image Understanding Workshop*, pages 1005–1010, 1993.

[50] C.C. Weems, C. Brown, J.A. Webb, T. Poggio, and J.R. Kender. Parallel processing in the DARPA Strategic Computing vision program. *IEEE Expert*, 6(5):23–38, October 1991.

[51] Y. Yacoob and L.S. Davis. Computational ground and airborne localization over rough terrain. In *Conf. on Computer Vision and Pattern Recognition*, pages 781–783, 1992.